

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the Application of: ,

HO SANG SUNG, ET AL.

Application No.:

Filed:

For: **APPARATUS FOR CODING
WIDE-BAND LOW BIT RATE
SPEECH SIGNAL**

Art Group:

Examiner:

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

REQUEST FOR PRIORITY

Sir:

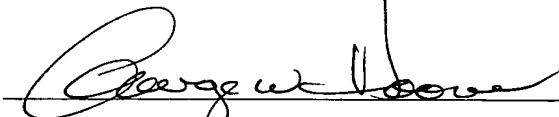
Applicant respectfully requests a convention priority for the above-captioned application, namely:

COUNTRY	APPLICATION NUMBER	DATE OF FILING
Korea	2003-15683	13 March 2003

☒ A certified copy of the document is being submitted herewith.

Respectfully submitted,

Blakely, Sokoloff, Taylor & Zafman LLP



George W Hoover, Reg. No. 32,992

Dated: December 30, 2003

12400 Wilshire Boulevard, 7th Floor
Los Angeles, CA 90025
Telephone: (310) 207-3800



별첨 사본은 아래 출원의 원본과 동일함을 증명함.

This is to certify that the following application annexed hereto
is a true copy from the records of the Korean Intellectual
Property Office.

출 원 번 호 : 10-2003-0015683
Application Number

출 원 년 월 일 : 2003년 03월 13일
Date of Application MAR 13, 2003

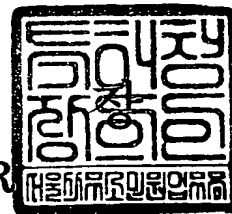
출 원 인 : 한국전자통신연구원
Applicant(s) Electronics and Telecommunications Research Inst



2003 년 11 월 28 일

특 허 청

COMMISSIONER



【서지사항】

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【참조번호】	0007
【제출일자】	2003.03.13
【국제특허분류】	G11B
【발명의 명칭】	광대역 저전송률 음성 신호의 부호화기
【발명의 영문명칭】	Apparatus for coding wide-band low bit rate speech signal
【출원인】	
【명칭】	한국전자통신연구원
【출원인코드】	3-1998-007763-8
【대리인】	
【성명】	이영필
【대리인코드】	9-1998-000334-6
【포괄위임등록번호】	2001-038378-6
【대리인】	
【성명】	이해영
【대리인코드】	9-1999-000227-4
【포괄위임등록번호】	2001-038396-8
【발명자】	
【성명의 국문표기】	성호상
【성명의 영문표기】	SUNG, Ho Sang
【주민등록번호】	710501-1822714
【우편번호】	305-325
【주소】	대전광역시 유성구 노은동 열매마을아파트 901동 401호
【국적】	KR
【발명자】	
【성명의 국문표기】	황대환
【성명의 영문표기】	HWANG, Dae Hwan
【주민등록번호】	630716-1010012

【우편번호】	305-729		
【주소】	대전광역시 유성구 전민동 청구나라아파트 103동 905호		
【국적】	KR		
【심사청구】	청구		
【취지】	특허법 제42조의 규정에 의한 출원, 특허법 제60조의 규정에 의한 출원심사를 청구합니다. 대리인 이영필 (인) 대리인 이해영 (인)		
【수수료】			
【기본출원료】	20	면	29,000 원
【가산출원료】	5	면	5,000 원
【우선권주장료】	0	건	0 원
【심사청구료】	7	항	333,000 원
【합계】	367,000 원		
【감면사유】	정부출연연구기관		
【감면후 수수료】	183,500 원		
【기술이전】			
【기술양도】	희망		
【실시권 허여】	희망		
【기술지도】	희망		
【첨부서류】	1. 요약서·명세서(도면)_1통		

【요약서】**【요약】**

본 발명은 음성신호 처리에 관한 것으로 특히 광대역 음성 신호의 부호화기에 관한 것이며 보다 상세하게는 광대역 저전송률 음성 신호의 부호화기에 관한 것으로, 고주파수로 샘플링된 음성 신호 프레임을 저주파수로 다운 샘플링 한 후, DC성분이 제거된 음성 신호 프레임을 생성하는 전처리 및 다운샘플링 블록; 상기 다운 샘플링된 음성 신호 프레임을 입력받아 상기 프레임의 선형 예측 계수를 구하고 이 계수를 ISP로 변환하여 양자화 시키며 상기 ISP의 인덱스를 생성하는 LPC분석 및 ISP양자화 블록; 상기 다운 샘플링된 음성 신호의 합성 필터의 여기 신호를 모델링하기 위한 잔차 신호를 산출하는 잔차신호 계산 블록; 상기 여기 신호의 모델링을 위한 랜덤벡터를 발생시키는 랜덤벡터 발생 블록; 상기 랜덤벡터를 스케일링하기 위한 이득을 산출하는 이득계산 블록; 및 상기 이득을 양자화하고, 상기 이득의 인덱스를 생성하는 이득 양자화 블록을 포함한다.

【대표도】

도 5

【명세서】**【발명의 명칭】**

광대역 저전송률 음성 신호의 부호화기{Apparatus for coding wide-band low bit rate speech signal}

【도면의 간단한 설명】

도 1은 일반적인 광대역 음성 신호 코덱의 오디오부의 기능적 구성도이다.

도 2는 16bit 선형 PCM 신호의 포맷을 나타낸 도면이다.

도 3은 일반적인 CELP 방식의 부호화기의 구성도이다.

도 4는 일반적인 CELP 방식의 복호화기의 구성도이다.

도 5는 본 발명이 제공하는 부호화기의 일실시예의 구성도이다.

도 6은 복호화기의 구성도이다.

도 7은 본 발명에 의한 부호화기의 비트 할당을 나타내는 도면이다.

도 8은 seed 생성의 방식을 C 프로그래밍 언어로 작성한 것을 보여주는 도면이다.

도 9는 이득의 양자화를 나타낸 도면이다.

<도면의 주요 부호에 대한 설명>

501: 전처리 및 다운샘플링(pre-processing & down sampling) 블록

502: LPC분석 및 ISP양자화(LPC analysis & ISP quantization) 블록

503: 잔차신호 계산(residual signal computation) 블록

504: 씨드 발생(seed generation) 블록

505: 랜덤벡터 발생(random vector generation) 블록

506: 이득 계산(gain computation) 블록

507: 갱신필터 메모리(update filter memory) 블록

508: 이득 양자화(gain quantization) 블록

【발명의 상세한 설명】

【발명의 목적】

【발명이 속하는 기술분야 및 그 분야의 종래기술】

- <19> 본 발명은 음성신호 처리에 관한 것으로 특히 광대역 음성 신호의 부호화기에 관한 것이며 보다 상세하게는 광대역 저전송률 음성 신호의 부호화기에 관한 것이다.
- <20> 음성신호의 부호화는 일반적으로 신호의 대역폭을 기준으로 하여 협대역 신호와 광대역 신호로 나누어 실시한다. 협대역 신호는 아날로그 입력 음성신호를 8 kHz로 샘플링을 한 16비트 선형 PCM 데이터를 음성 부호화기의 입력신호로 사용하며 광대역 신호는 아날로그 입력신호를 16 kHz로 샘플링 한 16비트 선형 PCM 데이터를 음성 부호화기의 입력신호로 사용한다.
- <21> 전자의 입력을 이용하는 신호를 대상으로 한 음성 부호화 기술들은 ITU-T의 표준안 중에서 G.711 ~ G.712의 PCM 방식과 G.720 ~ G.729 시리즈 같은 PCM 이외의 방법으로 압축하는 방식이 있으며, 후자의 입력을 대상으로 하는 음성 코덱은 ITU-T의 G.722와 G.722.1, 그리고 IMT-2000에 사용 예정인 AMR-WB(G.722.2)가 있다. 대표적인 음성 신호의 부호화 방식을 언급하면 다음과 같다.
- <22> 대표적인 협대역 음성의 부호화 방식의 하나로 ITU-T G.723.1을 들 수 있는데 이는 멀티미디어 신호를 저속으로 압축하기 위해 입력 음성을 5.3 또는 6.3 kbps의 dual rate로 압축하

여 복원하는 알고리즘이며 유선망의 음성 품질(toll quality)을 제공한다. 그리고 G.723.1은 과형 코딩과 파라미터 코딩이 혼합된 하이브리드 코딩 기술을 사용하며 CELP(Code Excited Linear Prediction) 계열의 음성 부호화 방식이다.

- <23> ITU-T G.722는 광대역 음성 신호의 부호화 방식이며 64,56,48 kbps의 전송률을 가진다. face-to-face 통신 품질을 보유하고 있으며 대역을 두 개의 서브밴드로 나눈 후 각 밴드를 ADPCM(Adaptive Differential Pulse Code Modulation)을 이용하여 부호화한다.
- <24> 3GPP AMR-WB(G.722.2)도 광대역 음성 신호의 부호화 방식이며 가장 최근에 표준화된 부호화 방식다. 이것은 늘어나는 이동통신의 수요를 확장하기 위해 IMT-2000에 응용할 목적으로 표준화되었다. 동일한 코덱을 ITU-T에서는 G.722.2로 명명하였으며 이는 유무선에 동시에 사용될 목적으로 표준화되었다. G.722.2은 9개의 전송률을 가지고 있으며 최대 23.85 kbps 이다. 최대 전송률에서 ITU-T G.722 64 kbps를 능가하는 음질을 제공한다.
- <25> 유선망 수준의 톨 음질(toll quality)을 제공하는 저속의 음성 부호화기는 높은 주파수 효율로 인해 이동통신 및 인터넷 텔레포니 등에서 새로운 서비스를 가능하게 하였다. 특히 인터넷 망에서의 VoIP(Voice over Internet Protocol)는 저렴한 전화요금으로 인해 기하급수적으로 확산이 되고 있는 상태이나 그리 좋은 평가는 받고 있지 못하다.
- <26> 인터넷을 이용한 음성 통신의 확산 추세에 걸림돌로 작용하는 낮은 음성 품질과 음성 처리 지연으로 인해 생기는 서비스 저하를 다양한 방법으로 극복해 보려는 시도가 계속되고 있다. 이런 시도의 일환으로 광대역 음성신호를 대상으로 VoIP 서비스를 시도하고 있는데 광대역 신호를 대상으로 한 서비스는 음성 품질 측면에서 많은 성능 향상을 가져오고 있다.

<27> 가장 최근에 발표된 광대역 음성 코덱(codec)은 위에서 언급한 바와 같이 AMR-WB인데 일반적인 CELP(Code Excited Linear Prediction) 방식을 사용하고 있는데 9개의 전송률 모드가 있으며 가장 낮은 전송률이 6.6 kbps이다. 이 음성 코덱의 단점은 소스 제어 가변 전송 레이트(source controlled variable rate)를 지원하고 있지 않다는 즉, 규격에서 정한 레이트만 사용하기 때문에 입력 음성 신호의 특성이 반영되지 않는다는 점이다. 그리고 규격에서 제공하는 VAD(Voice Activity Detection) 알고리즘도 단지 입력 신호가 유성음인지 무성음인지(voiced or unvoiced)만 판단하기 때문에 특히 묵음(silence)으로 판별된 신호 구간의 전송에 있어서 문제점이 있다.

<28> 따라서 소스 제어 가변 전송 레이트를 완벽하게 지원하려면 입력 신호를 그 특성에 따라 정확하게 분류할 수 있는 VAD 알고리즘이 필요하며 이 결과에 따라 각 입력의 특성에 따른 전송률을 유연하게 결정함이 필요하다.

【발명이 이루고자 하는 기술적 과제】

<29> 따라서 본 발명은 상기의 문제점을 해결하고 그 필요성에 부응하기 위해 창안된 것으로 본 발명의 목적 및 이루고자 하는 기술적 과제는 음성 신호의 특성에 따른 전송률을 유연하게 결정할 수 있는 특히, VAD 알고리즘에 의해 묵음(silence)로 판별된 음성 신호 구간을 위한 광대역 저전송률 음성 신호 부호화기를 제공함에 있다.

【발명의 구성 및 작용】

<30> 이와 같은 목적 및 기술적 과제를 달성하기 위하여 본 발명이 제공하는 광대역 저전송률 음성 코덱 부호화기는 고주파수로 샘플링된 음성 신호 프레임을 저주파수로 다운 샘플링 한 후, DC성분이 제거된 음성 신호 프레임을 생성하는 전처리 및 다운샘플링 블록; 상기 다운 샘플링

플링된 음성 신호 프레임을 입력받아 상기 프레임의 선형 예측 계수를 구하고 이 계수를 ISP로 변환하여 양자화 시키며 상기 ISP의 인덱스를 생성하는 LPC분석 & ISP양자화 블록; 상기 다른 샘플링된 음성 신호의 합성 필터의 여기 신호를 모델링하기 위한 잔차 신호를 산출하는 잔차신호 계산 블록; 상기 여기 신호의 모델링을 위한 랜덤벡터를 발생시키는 랜덤벡터 발생 블록; 상기 랜덤벡터를 스케일링하기 위한 이득을 산출하는 이득 계산 블록; 및 상기 이득을 양자화하고, 상기 이득의 인덱스를 생성하는 이득 양자화 블록을 포함함을 그 특징으로 한다.

<31> 우선, 이해의 편의를 위해 본 발명의 구현 방식을 간단히 언급하면 다음과 같다.

<32> 본 발명은 광대역 음성신호를 하위대역(50~6400Hz)과 상위대역(6400~7000Hz)으로 분할하여, 이중에서 하위대역에 해당하는 50~6400Hz대역의 신호를 1.0kbps의 저전송률로 부호화/복호화하는 방식에 대한 것이다.

<33> 본 발명에서 제안한 부호화/복호화 방식은 VAD 알고리즘에 의해 음성 신호를 그 특성에 따라 voiced, unvoiced, music, background noise, onset, silence 등으로 구분하는 경우 이중에서 묵음(silence)에 해당하는 구간을 위한 저전송률 음성 부호화기 알고리즘을 제안하는 데 그 목적이 있다. 여기서 묵음으로 선택되는 구간의 신호는 낮은 레벨의 잡음(noise) 신호와 음성의 묵음신호를 의미한다.

<34> 본 발명의 구현에 적용되는 기본적인 방식은 LP(Linear Prediction) 분석을 이용하는 CELP(Code Excited Linear Prediction) 방식이다.

<35> 본 발명의 구현 태양을 간단히 살펴보면 먼저 음성신호를 20ms의 프레임으로 분리한다. 이러한 20ms 프레임에 대해서 단구간 상관도를 나타내는 LPC 계수(Linear Prediction Coding coefficient)를 구하는데 구하는 과정에서 선형 예측을 위해 5ms의 lookahead를 이용하므로 총

지연 시간은 25ms가 된다. 사용되는 LPC 계수의 차수는 16차이고 이 계수를 수학적으로 등가 이면서 양자화와 안정화 체크에 유리한 ISP(Immittance Spectral Pair) 계수로 변환한다.

<36> ISP 계수는 분할 양자화되며 분할 양자화에는 총 14 비트가 할당된다. 양자화된 LPC 계수는 2번째 부프레임을 위한 계수이며 첫번째 부프레임을 위한 계수는 이전 프레임에서 구한 LPC 계수와의 보간(interpolation)을 통하여 구할 수 있다. 각 부프레임에 대해서 양자화된 LPC 계수를 이용하여 분석 필터를 구성한 후 입력 신호를 통과시키면 잔차신호가 발생한다. 이 잔차신호를 모델링하기 위해 본 발명에서는 랜덤 시퀀스(random sequence)를 발생시켜서 적당한 이득을 곱하는 방식을 사용한다. 이득은 잔차신호와 랜덤 시퀀스 간의 크로스 상관(cross correlation)을 이용하여 구한다. 구해진 이득은 2차 MA예측기와 스칼라 양자화기를 사용하여 양자화된다. 이득의 양자화에는 부프레임당 3 비트, 총 6 비트를 이용하여 양자화한다. 그리고 다음 프레임을 위해서 메모리를 업데이트 한다.

<37> 이하 본 발명의 구성 및 작용을 본 발명의 최적의 실시예에 근거, 첨부 도면을 참조하여 상세히 설명하되 도면의 구성요소들에 참조번호를 부여함에 있어서 동일 구성요소에 대해서는 비록 다른 도면상에 있더라도 동일 참조번호를 부여하였으며 당해 도면에 대한 설명시 필요한 경우 다른 도면의 구성요소를 인용할 수 있음을 미리 밝혀둔다.

<38> 도 1은 일반적인 광대역 음성 신호 코덱의 오디오부의 기능적 구성도이다.

<39> 아날로그 음성 입력신호는 ADC/DAC(10)를 거쳐 디지털화되며 디지털화 된 음성 입력신호는 광대역 음성 코덱(Wideband speech codec, 11)으로 입력된다. 부호화/복호화기(12)는 입력 신호를 부호화하여 패킷화한 후 채널(channel, 13)로 전송한다. 채널(13)을 통해 전송 받은 패킷데이터는 부호화/복호화기(12)에 의해 복호화되며 복호화된 음성신호는 ADC/DAC(10)에 의해 아날로그화 된 후 스피커로 아날로그 음성 신호로 출력된다.

- <40> ADC/DAC(10)를 거쳐 광대역 음성 코덱(Wideband speech codec, 11)으로 입력되는 신호의 형식은 16bit 형식을 가지는 16bit 선형 PCM(16bit linear Pulse Code Modulation) 신호이다. 이 입력신호의 구체적인 형태는 도 2에 제시되어 있다. 도 2에서 하위 2비트는 0으로 채워지므로 코덱에서 신호를 처리할 때는 비트를 2비트를 오른쪽으로 쉬프트해야 한다.
- <41> 도 1에 제시된 광대역 음성 코덱(11)을 낮은 전송률로 구현하기 위해 일반적으로 CELP(Code Excited Linear Prediction) 방식의 부호화기를 많이 이용한다. 일반적인 CELP 방식의 부호화기는 도 3에 제시되어 있다.
- <42> 먼저 입력 음성 신호[s(n)]는 전처리(301)에 의해 전처리 과정을 거친 후 LPC 분석(302)을 거치게 된다.
- <43> **【수학식 1】**
$$A(z) = 1 + \sum_{i=1}^m a_i z^{-i}$$
- <44> 여기서 $A(z)$ 은 분석필터를 의미하며 LPC 분석(302)을 통해서 구할 수 있다. $\{a_i\}$ 은 선형예측계수(Linear Prediction Coefficient : LPC)을 의미한다. 분석된 선형예측계수 $\{a_i\}$ 은 양자화 과정을 거친 후 LP 합성필터(303)을 구성하며 이 필터의 함수는 아래 수학식 2에 제시되어 있다. m 의 값에 따라 예측 차수가 결정이 되는데 보통 협대역의 음성코덱은 10을 차수로 사용하며 광대역의 음성 코덱은 16~20 정도의 차수를 사용한다.
- <45> **【수학식 2】**
$$H(z) = \frac{1}{\hat{A}(z)} = \frac{1}{1 + \sum_{i=1}^m \hat{a}_i z^{-i}}$$
- <46> 수학식 2에서 $H(z)$ 은 LPC 합성필터(303)를 의미하며 $\hat{A}(z)$ 은 양자화된 $A(z)$ 이고 \hat{a}_i 는 양자화된 선형예측계수(LPC)이다. LPC는 전송을 위해서 양자화가 되며 양자화된 계수는 LPC 합성필터를 구성한다. LPC 합성필터(303)를 이용하여 페루프를 돌면서 여기 신호(excitation

signal)를 구한다. 여기신호를 구하기 위한 대상신호는 보통 입력신호를 적응가중 필터(304)를 사용한 결과를 이용하여 구할 수 있다. 이와 같이 적응가중 필터(304)를 사용한 신호를 대상으로 여기신호를 분석하여 구하면 더 우수한 복원음성을 얻을 수 있다. 여기 신호로는 적응 코드북(309)에서 구해지는 장구간 상관도 신호와 고정 코드북(307)에서 구해지는 단구간 상관도 신호가 있다. 이 두 신호에 적절한 이득(G_c , G_p)을 곱한 후에 합하면 LPC 합성필터(303)의 여기 신호가 된다.

<47> CELP방식은 고정 코드북(307)과 적응 코드북(309)의 탐색시 직접 합성을 한 후에 분석을 하는 AbS(Analysis by Synthesis) 방식을 사용한다. 직접 합성을 해야 하므로 많은 계산량을 필요로 한다. 수학식 3은 적응 코드북(309)에서 구해지는 장구간 상관도 신호에 대한 LPC 합성 필터(303)의 함수를 나타낸다.

<48> **【수학식 3】**
$$\frac{1}{B(z)} = \frac{1}{1 - G_p z^{-T}}$$

<49> 수학식 3에서 중요한 값은 피치 분석(305)을 통해 얻어진 피치(pitch) 주기 T와 적절한 이득 G_p 이다. 주기 T 이전의 과거의 합성신호[$\{z\}^{-T}$]를 이용하여 현재 신호를 장구간으로 예측을 하며 여기에 이득(G_p)을 곱해주면 현재의 장구간 상관도 신호[$B(z)$]가 된다. 이 장구간 상관도 신호의 중요 계수인 T와 G_p 를 구한 후 더 정밀한 여기 신호를 구하기 위해 고정 코드북 탐색(306)을 실시한다.

<50> 고정 코드북 탐색(306)의 대상신호는 장구간 상관도를 구하기 위한 대상 신호에서 구해진 장구간 상관도 신호를 제하고 남은 신호이다. 고정 코드북(307)은 다양한 방식으로 구현이 되며, 최근에 가장 많이 사용되는 방식이 대수적 코드북(algebraic codebook) 구조이다. 이 방식은 코드북을 저장하기 위한 메모리가 필요 없으며 고속으로 원하는 innovation 신호를 구할

수 있다. 이 구조의 단점은 계산량이 많다는 것인데 최근에 다양한 빠른 알고리즘이 제안되어 있으므로 계산량이 그리 문제되지는 않는다. 대수적 코드북탐색을 통해서 나오는 계수는 펄스의 위치정보와 부호정보이다. 이렇게 고정 코드북이 구해지면 이 코드북에 해당하는 이득을 구해야 한다. 고정 코드북의 이득은 적응코드북의 이득과 함께 페루프로 구해진다. 구해진 이득들은 이득 양자화 블록(311)에서 벡터 양자화 된다. 이와같이 전체 프레임에 대해 분석이 끝나면 구해진 계수들을 이용하여 파라미터 인코딩 블록(312)에서 비트스트림으로 부호화해서 전송을 한다.

<51> 일반적인 CELP 방식의 복호화기는 도 4에 나타나 있다. 복호화기는 도 3의 부호화기에서 전송된 비트 스트림을 이용하여 파라미터 디코딩 블록(401)에서 각 계수로 변환시킨 후에 각각 해당하는 모듈에서 사용한다. 먼저 복호화된 LPC계수를 이용하여 LPC 합성필터(406)을 구성하고 고정 코드북(402)과 적응 코드북(404)의 인덱스를 복호화하여 각각 이득(G_c, G_p)을 곱한 후 여기 신호를 만들어 낸다. 이 신호를 LPC 합성필터(406)를 통과시키면 합성 신호가 발생하며 이 신호를 고품질의 아날로그 출력 음성으로 만들기 위해서 후처리 필터(407)를 거친다.

<52> 지금까지 일반적인 CELP의 구조에 대해 설명하였다. 본 발명은 이와 같은 CELP 구조를 이용하지만 저 전송률을 위해 피치 분석(305)과 고정 코드북 탐색(306)을 하지 않고, 이 부분에 랜덤 시퀀스를 발생시켜 여기 신호를 모델링한다.

<53> 도 5는 본 발명이 제공하는 부호화기의 구성도이다. 본 발명이 제공하는 음성 코덱 부호화기(50)는 50~6400Hz의 대역을 사용하며 1.0 kbps의 전송률을 갖도록 설계되었다. ISP, 이득(Gain)의 2가지 특징 파라미터(ISP Index, Gain Index)를 추출하여 복호화기에 전송하게 되며 각 파라미터 당 2개의 부프레임으로 구성되고 각 프레임에 대한 비트 할당은 도 7에 제시되어 있다.

- <54> 본 발명이 제공하는 도 5의 부호화기에서는 프레임 단위로 분석이 이루어지게 된다.
- <55> 전처리 및 다운샘플링(pre-processing & down sampling) 블록(501)에서는 16kHz로 샘플링된 입력음성신호를 12.8kHz로 다운 샘플링 한 후, 50Hz이하의 DC성분이 제거된 신호를 만든다.
- <56> LPC분석 및 ISP양자화(LPC analysis & ISP Quantization) 블록(502)은 이 신호를 입력받아 자기 상관 함수(autocorrelation function)를 통한 Levinson-Durbin 방법을 이용하여 선형 예측 계수를 구한다. 선형 예측의 차수는 16차이며 이 계수를 이용하여 앞에서 언급한 수학식 1에 의해 음성 신호의 단구간 상관도[A(z)]를 분석한다.
- <57> 상기 a_i 을 양자화시켜 \hat{a}_i 을 획득한 후, \hat{a}_i 을 이용하여 합성 필터를 구성하므로 이 계수를 이용하여 양자화 에러를 최소화하면서 양자화하는 과정이 중요하다. 하지만 LPC 계수는 동적 범위가 크기 때문에 양자화에 불리하다. 그래서 동적 범위도 적고 안정성 체크도 쉬우면서 수학적으로 LPC 계수와 등가인 ISP(Immittance spectral pairs) 계수로 변환하여 양자화를 한다.
- <58> ISP 계수의 양자화는 분할 벡터 양자화(Split Vector Quantization, SVQ)방식으로 양자화된다. 양자화에는 총 14비트가 소요되며 2 split으로 이루어져 있다. 두 개의 split 된 코드북을 이용하여 각각 7비트로 양자화를 한다.
- <59> 양자화된 단구간 상관도를 이용한 합성 필터는 앞에서 언급한 수학식 2과 같이 나타난다. 수학식 2에서 \hat{a}_i 는 양자화된 LPC 계수를 의미하고 m은 예측 차수를 의미하며 본 발명에서는 m값으로 16을 사용한다.

- <60> 나머지 과정은 앞에서 구해진 LP 합성필터의 여기 신호를 모델링하는 과정이며 부프레임 단위로 수행된다.
- <61> 먼저 잔차신호 계산 블록(503)은 전처리 및 다운샘플링 블록(501)의 출력 신호를 수학적식 3에 의한 분석 필터에 통과시켜 잔차(residual) 신호를 구한다. 이 잔차 신호는 합성 필터의 여기 신호를 모델링하는 타겟(target) 신호가 된다.
- <62> 이 여기 신호를 모델링하기 위해 랜덤벡터를 이용한다. 이 랜덤벡터는 통상 가우시안(Gaussian)이다. 모델링은 가우시안 랜덤벡터를 이용하여 랜덤 시퀀스를 발생시킨 후 여기에 적절한 이득(gain)을 곱하는 방식을 사용한다. 랜덤벡터 발생 블록(505)을 통해 구해지는 랜덤 벡터는 씨드 발생(seed generation) 블록(504)을 통해 아래 수학적식 4를 이용하여 구해지는 seed를 입력받아 도 7에 제시된 부프레임의 샘플 개수만큼 저장하는 방식을 이용하여 구할 수 있다. seed는 계속 업데이트 되므로 한번만 결정해 주면 나머지는 순서대로 생성이 된다.
- <63> 【수학적식 4】 $seed = (\text{Word16})(seed * 31821(=0x7c4d) + 13849(=0x3619))$
- <64> (Word16)은 16비트 integer 값을 의미한다. seed는 항상 위 식에 의해 업데이트 되지만 만일 프레임 삭제(frame erasure)가 발생하면 인코더와 디코더가 다른 값을 갖는 현상이 발생한다. 이런 프레임 삭제에 대처하기 위해 seed값을 전송된 파라미터를 이용하여 생성시키는 방식을 이용한다.
- <65> 씨드 발생 블록(504)을 통한 seed 생성은 LPC분석 및 ISP양자화 블록(502)으로부터 전송된 인덱스 2개를 이용하여 도 8에 제시된 방식으로 구할 수 있다.
- <66> 도 8은 seed 생성의 방식을 C 프로그래밍 언어로 작성한 것을 보여주는 도면이다.

- <67> 먼저 ①에서 lpc_ind[0]은 전송된 LPC 파라미터의 첫번째 인덱스를 의미한다. ②에서 lpc_ind[1]은 전송된 LPC 파라미터의 두번째 인덱스를 의미한다.
- <68> seed 0을 구하기 위해 ③에서 lpc_ind[0]를 shift left 8bit를 한 후, ④에서 이 값과 lpc_ind[1]을 배타적 OR연산을 한 결과를 16비트의 자연수로 저장한다. 그리고 seed 1을 구하기 위해 ⑤에서 lpc_ind[1]을 8비트 왼쪽으로 쉬프트한 후, ⑥에서 이 값과 lpc_ind[0]를 배타적 OR연산을 한 결과를 16비트의 자연수로 저장한다. 이렇게 해서 seed 0과 seed 1이 결정되면 ⑦, ⑧에서 이중 큰 값을 최종적으로 seed로 결정한다.
- <69> 랜덤벡터 발생 블록(505)은 이렇게 구해진 seed를 이용하여 부프레임 단위로 랜덤벡터를 구하게 되며 각 서브프레임의 랜덤벡터는 128개가 된다.
- <70> 이득 계산(Gain computation) 블록(506)은 구해진 랜덤벡터에 곱해지는 이득을 구하는 부분이다. 즉, 최종적으로 이득에 의해서 스케일링된 랜덤벡터는 LP 합성필터(604)의 여기 신호가 된다.
- <71> 이득은 앞에서 구해진 LP 잔차 신호인 $r(n)$ 와 랜덤벡터를 이용하여 아래의 수학적 식 5에 의해 $gain(\{g\}_{\{s\}})$ 을 구한다.

<72>

$$g_s = 0.75 * \sqrt{\frac{\sum_{n=0}^{127} [r(n)]^2}{\sum_{n=0}^{127} [e_{rand}(n)]^2}}$$

【수학적 식 5】

- <73> 0.75는 이득 감쇄 팩터(gain attenuation factor)이며, $e_{rand}(n)$ 은 랜덤벡터를 의미한다. 이득 양자화(gain quantization) 블록(508)에서 현재 부프레임의 $gain[\{g\}_{\{s\}}(n)]$ 의 양자화는 2차 MA 예측기(Moving Average predictor, 91)를 이용하여 예측된 값을 뺀 예측에러벡터를

양자화함으로써 이루어진다. 도 9는 $\text{gain}[\{g\}_{\{s\}}(n)]$ 의 양자화기를 나타낸 것이다. 양자화기(90)의 입력신호인 예측에러벡터 $c(n)$ 은 다음과 같이 나타낼 수 있다.

<74> **【수학식 6】** $c(n) = g_s(n) - p(n)$

<75> 여기서 $g_s(n)$ 은 이득 계산 블록(506)에서 구해진 이득이며, 예측 벡터 $p(n)$ 은 수학식 7과 같이 이전 부프레임에서 이미 양자화된 예측에러 벡터 $\hat{c}(n)$ 을 이용한 2차 MA 예측기(91)를 사용해서 구한다.

<76> **【수학식 7】** $p(n) = \sum_{j=1}^2 g_j \hat{c}(n-j)$

<77> 여기서 $\hat{c}(n)$ 는 n번째 프레임에서 양자화된 예측 에러벡터이며 g_j 는 MA 예측기(91)의 계수이다. 본 발명에서 $[g_1, g_2]$ 값은 $[0.28, 0.11]$ 로 설정된다. 양자화된 $\text{gain}[\hat{g}_s(n)]$ 은 수학식 8과 같이 양자화된 예측에러 벡터 $[\hat{c}(n)]$ 에 예측벡터 $[p(n)]$ 를 더함으로써 구할 수 있다.

<78> **【수학식 8】** $\hat{g}_s(n) = \hat{c}(n) + p(n)$

<79> 도 9의 양자화기는 현재 프레임의 예측 에러벡터인 $c(n)$ 값을 스칼라 양자화한다. 스칼라 양자화는 총 3비트를 사용하므로 8개의 코드워드를 갖는다. 양자화 과정이 끝나면 갱신필터 메모리(Update filter memory) 블록(507)에서 다음 프레임을 위한 메모리 갱신을 실시한다.

<80> 메모리 갱신은 부호화 대상 음성신호버퍼, 가중 처리된 음성신호 버퍼, 그리고 여기 신호 버퍼를 갱신한다. 프레임별로 부호화가 끝난 후, 디코더로 실제 전송되는 인덱스는 LPC 양자화기의 인덱스(LPC Index) 14 비트와 이득 양자화기의 인덱스(Gain index) 6 비트로 총 20 비트가 된다.

<81> 도 6은 복호화기(decoder)의 구성도이다. 복호화기는 전송된 인덱스(LPC Index, Gain index)를 이용하여 LP 합성 필터(604)를 구성하고 603 블록의 이득($\{g\}_{\{s\}}$)을 구한다. 그리고 전송된 LPC 인덱스를 이용하여 씨드 발생 블록(601)을 통해 도 8에 제시된 방식과 마찬가지로 seed를 구하고, 이 seed를 이용하여 랜덤벡터 발생 블록(602)은 랜덤벡터를 생성한다. 이 랜덤벡터에 이득(g_s)을 곱한 신호가 LP 합성 필터(Synthesys Filter)(604)의 여기 신호가 된다. 이 여기 신호를 LP 합성 필터(604)를 통과시키면 합성된 음성신호가 복원된다.

<82> 이제까지 본 발명에 대하여 그 바람직한 실시 태양을 중심으로 살펴보았다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다.

<83> 그러므로 개시된 실시예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

【발명의 효과】

<84> 본 발명에 의하면, 음성 신호의 특성에 따른 전송률을 유연하게 결정할 수 있으며 특히, 묵음(silence)으로 판별된 음성 신호 구간을 위한 광대역 저전송률 음성 신호를 효율적으로 부호화/복호화할 수 있다. 아울러 본 발명을 이용한 후 보다 높은 밴드 모델링(higher band modeling) 기법을 이용하여 6.4 ~ 7kHz의 대역만 발생시켜서 더해주면 완전한 광대역 음성 부호화를 실현할 수 있는 이점도 있다.

【특허청구범위】**【청구항 1】**

고주파수로 샘플링된 음성 신호 프레임을 저주파수로 다운 샘플링 한 후, DC성분이 제거된 음성 신호 프레임을 생성하는 전처리 및 다운샘플링 블록;

상기 다운 샘플링된 음성 신호 프레임을 입력받아 상기 프레임의 선형 예측 계수를 구하고 이 계수를 ISP로 변환하여 양자화 시키며 상기 ISP의 인덱스를 생성하는 LPC분석 및 ISP 양자화 블록;

상기 다운 샘플링된 음성 신호의 합성 필터의 여기 신호를 모델링하기 위한 잔차 신호를 산출하는 잔차신호 계산 블록;

상기 여기 신호의 모델링을 위한 랜덤벡터를 발생시키는 랜덤벡터 발생 블록;

상기 랜덤벡터를 스케일링하기 위한 이득을 산출하는 이득 계산 블록; 및

상기 이득을 양자화하고, 상기 이득의 인덱스를 생성하는 이득 양자화 블록을 포함함을 특징으로 하는 광대역 저전송률 음성 신호의 부호화기.

【청구항 2】

제 1 항에 있어서, 상기 모델링은

상기 음성 신호 프레임의 두 개 부프레임 단위로 이루어지며, 상기 랜덤벡터를 이용하여 랜덤 시퀀스를 발생시킨 후 상기 이득을 곱하여 이루어짐을 특징으로 하는 광대역 저전송률 음성 신호의 부호화기.

【청구항 3】

제 2 항에 있어서, 상기 랜덤벡터는

소정의 방식으로 생성된 seed를 상기 부프레임의 샘플 개수만큼 저장하여 발생됨을 특징으로 하는 광대역 저전송률 음성 신호의 부호화기.

【청구항 4】

제 3 항에 있어서, 상기 seed의 생성은

상기 LPC분석 & ISP양자화 블록으로부터 전송된 2개의 인덱스를 이용하며, 상기 인덱스 중 첫 번째 인덱스를 8 비트 왼쪽으로 쉬프트한 값을 구한 후 이 값과 상기 인덱스 중 두 번째 인덱스의 배타적 OR 연산 결과를 제1 seed값(seed 0)로 하고, 상기 두 번째 인덱스를 8 비트 왼쪽으로 쉬프트한 값을 구한 후 이 값과 상기 첫 번째 인덱스의 배타적 OR 연산을 한 결과를 제2 seed값(seed 1)으로 하여 상기 seed 0과 seed 1 중 큰 값을 최종 seed 값으로 결정하여 이루어짐을 특징으로 하는 광대역 저전송률 음성 신호의 부호화기.

【청구항 5】

제 1 항에 있어서, 상기 이득은

상기 잔차 신호와 상기 랜덤벡터를 근거로 산출됨을 특징으로 하는 광대역 저전송률 음성 신호의 부호화기.

【청구항 6】

제 1 항에 있어서, 상기 ISP 인덱스와 이득 인덱스는 각각 14 비트, 6 비트로 양자화됨을 특징으로 하는 광대역 저전송률 음성 신호의 부호화기.

【청구항 7】

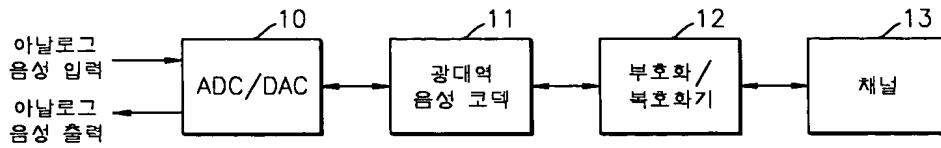
제 1 항에 있어서, 상기 이득의 양자화는

상기 이득에서 이전 프레임에 대하여 이미 양자화된 예측에러벡터 값에 대하여 예측된 값을 뺀 현재 예측에러벡터를 양자화하여 이루어짐을 특징으로 하는 광대역 저전송률 음성 신호의 부호화기.

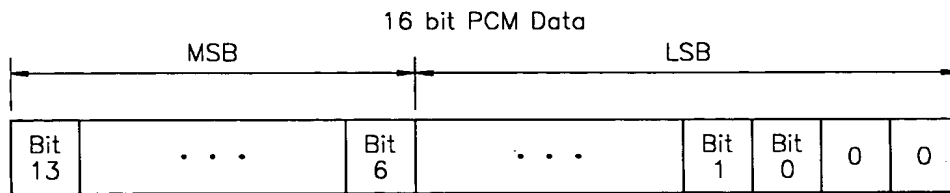


【도면】

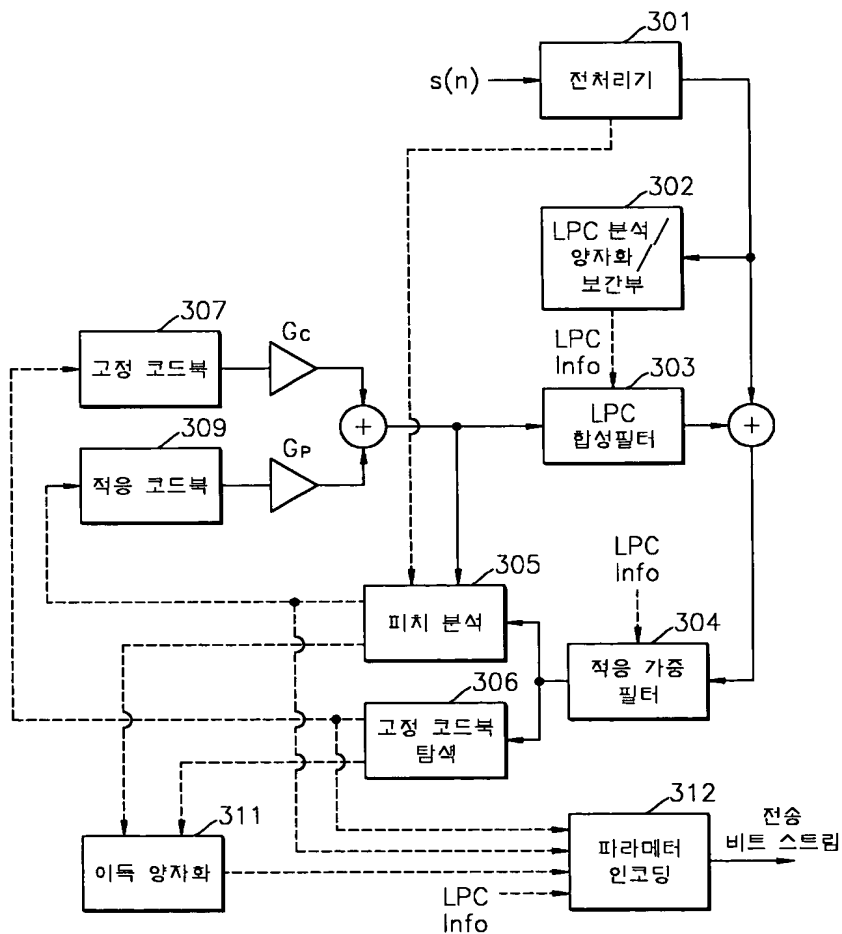
【도 1】



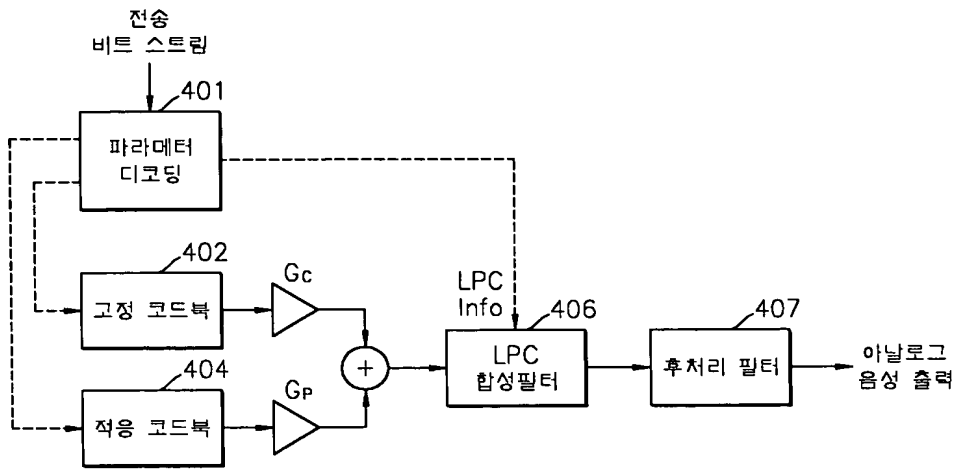
【도 2】



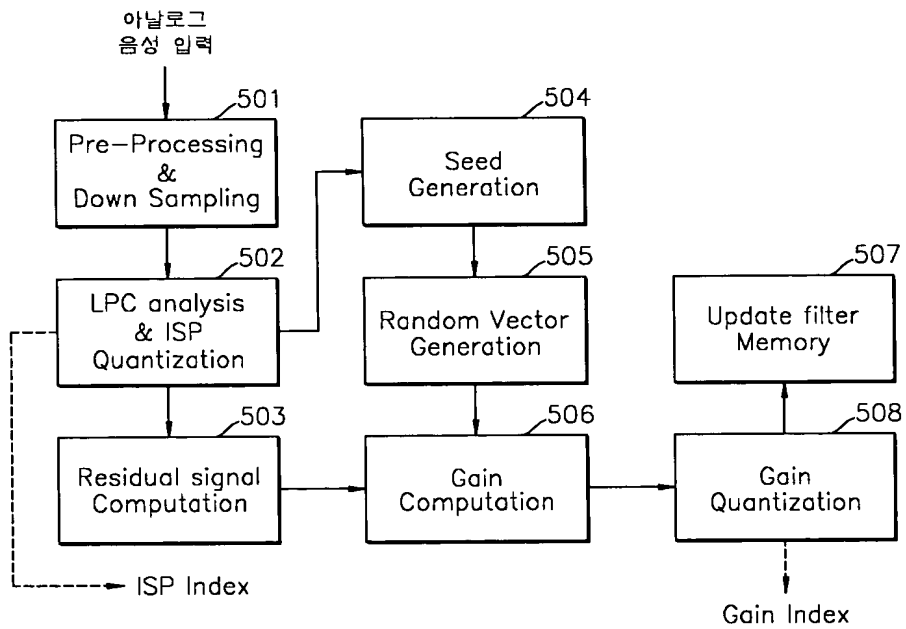
【도 3】



【도 4】



【도 5】

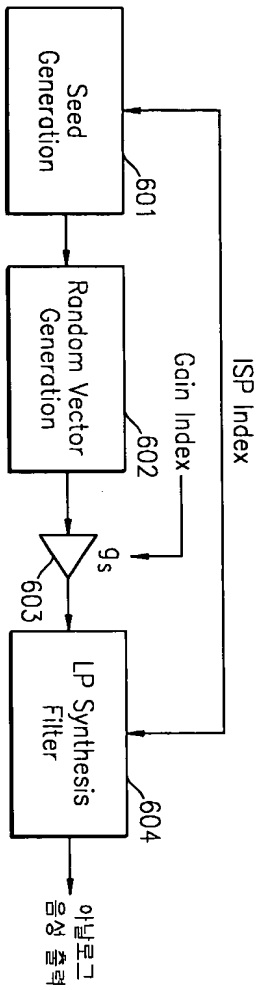




1020030015683

출력 일자: 2003/12/4

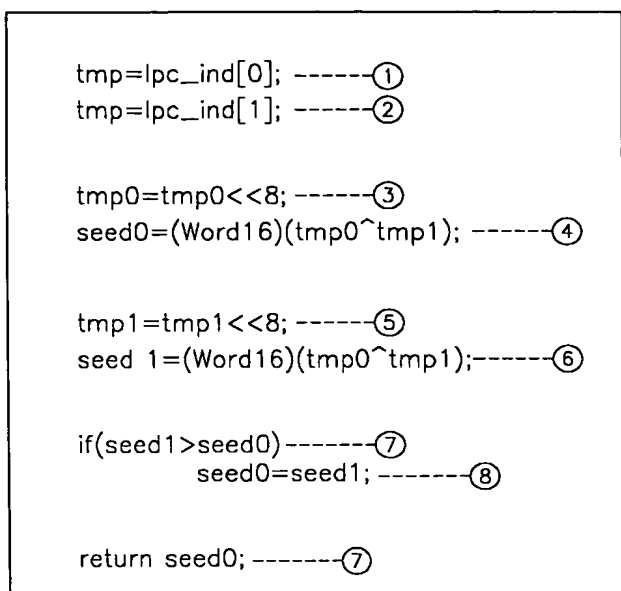
【도 6】



【도 7】

Rate	Parameter	1 st subframe	2 nd subframe	Total per frame
1.0 kbps	ISP			14 bit
	Gain	3	3	6 bit
	Total			20bit

【도 8】



【도 9】

